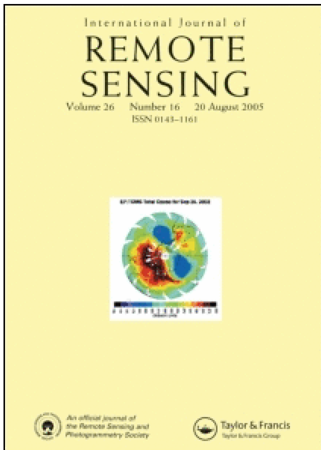

International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713722504

Spatial agreement between two land-cover data sets stratified by agricultural eco-regions

P. Y. Chen [a]; M. Di Luzio [a]; J. G. Arnold [b]
[a] Blackland Research and Extension Center, Texas Agricultural Experiment Station, Temple, TX 76502
[b] USDA-ARS, Grassland, Soil and Water Research Lab., Temple, TX 76502

Online Publication Date: 10 August 2006

To cite this Article: Chen, P. Y., Luzio, M. Di and Arnold, J. G. (2006) 'Spatial agreement between two land-cover data sets stratified by agricultural eco-regions', International Journal of Remote Sensing, 27:15, 3223 — 3238

To link to this article: DOI: 10.1080/01431160600567803
URL: http://dx.doi.org/10.1080/01431160600567803

International Journal of Remote Sensing
Vol. 27, No. 15, 10 August 2006, 3223–3238

Spatial agreement between two land-cover data sets stratified by agricultural eco-regions

P. Y. CHEN*†, M. DI LUZIO† and J. G. ARNOLD‡
†Blackland Research and Extension Center, Texas Agricultural Experiment Station, 720 East Blackland Road, Temple, TX 76502, USA
‡USDA-ARS, Grassland, Soil and Water Research Lab., 808 East Blackland Road, Temple, TX 76502, USA

(*Received 23 August 2005; in final form 4 January 2006*)

Two of the most widely used land-cover data sets for the United States are the National Land-Cover Data (NLCD) at 30-m resolution and the Global Land-Cover Characteristics (GLCC) at 1-km nominal resolution. Both data sets were produced around 1992 and expected to provide similar land-cover information. This study investigated the spatial distribution of NLCD within major GLCC classes at 1-km unit over a total of 11 agricultural-related eco-regions across the continental United States. Our results exhibited that data agreement or relationship between the GLCC and NLCD was higher for the eco-regions located in the corn belt plains with homogeneous or less complicated land-cover distributions. The GLCC cropland primarily corresponded to NLCD row crops, pasture/hay and small grains, and was occasionally related to NLCD forest, grassland and shrubland in the remaining eco-regions due to high land-cover diversity. The unique GLCC classes of woody savanna and savanna were mainly related to the NLCD orchard and grassland, respectively, in the eco-region located in the Central Valley of California. The GLCC urban/built-up among vegetated areas strongly agreed to the NLCD urban for the eco-regions in the corn belt plains. A set of sub-class land-cover information provided through this study is valuable to understand the degrees of spatial similarity for the major global vegetated classes. The sub-class information from this study provides reference for substituting less-detailed global data sets for detailed NLCD to support national environment studies.

1. Introduction

Land-cover information has become essential and crucial for environmental management and natural resource studies. The appropriate use of this information is usually a consequence of the derived spatial scale. Coarse spatial resolution (e.g. one to several kilometres) satellite sensors offer the advantage of frequent coverage of large areas, and the data plays an important role in many global environmental and climatic studies. Fine spatial resolution (e.g. one to a few metres) satellite sensors provide detailed land cover information on a local basis, and the data is generally applied to detail-oriented studies.

In the United States, an intermediate-scale land-cover data, National Land-Cover Data (NLCD), is required by an increasing number of applications at regional and national scales to support a wide range of management, monitoring and modelling

*Corresponding author. Email: pchen@brc.tamus.edu

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2006 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160600567803

activities in the fields of agriculture, forestry, hydrology and wildlife. The NLCD was developed at 30 m resolution using the 1992 LANDSAT Thematic Mapper (TM) data. The goal of NLCD was to provide a generalised, consistent and seamless land-cover data for the conterminous US (Vogelmann *et al*. 1998, Vogelmann *et al*. 2001). The NLCD has been continuously utilised to support national environmental assessments using watershed-based hydrologic models such as Soil and Water Assessment Tool (SWAT) (Arnold *et al*. 1998) and Spatially Referenced Regression on Watershed (SPARROW) (Smith *et al*. 1997). A general concern with NLCD is the update of the information content that takes extensive time and requires tremendous processing effort. In fact, the new version of national land-cover data based on the 2000 vintage LANDSAT data is still being completed by several federal agencies forming the Multi-Resolution Land Characteristics Consortium (MRLC) (Homer *et al*. 2002).

In order to meet the needs of multi-temporal and extended spatial land-cover data, several global land-cover data sets were generated at 1 km resolution over the last few years. The Joint Research Center (JRC) of the European Commission (EC) implemented the Satellite Pour l'Observation de la Terre (SPOT) VEGETATION satellite data to produce a global land-cover for the year 2000 (Giri *et al*. 2005). The National Aeronautics and Space Administration (NASA) in the US collaborated with Boston University to develop a global land-cover data using the Moderate Resolution Imaging Spectrometer (MODIS) satellite data (Friedl *et al*. 2002).

Previously, another global land-cover data set, Global Land-Cover Characteristics (GLCC), was produced at 1 km nominal spatial resolution for the entire world (Brown *et al*. 1993, Loveland *et al*. 2000). The GLCC was developed using Advanced Very High Resolution Radiometer (AVHRR) satellite data collected across 1992 through 1993, like for NLCD (Eidenshink and Faundeen 1994). The goal of GLCC was to produce flexible large-area land cover databases to meet evolving requirements of the earth science research community.

Other land-use related data sets such as eco-regions of the continental US provide multipurpose spatial frameworks allowing a wide range of environmental applications. In recent years, a growing number of scientists and public agencies have developed maps depicting eco-regions as analysis units for environmental studies (Loveland and Merchant 2004, Plummer 2000). Eco-regions can be defined at any scale as long as they represent areas with relative homogeneity in ecosystems (Omernik and Bailey 1997). The US Environmental Protection Agency (EPA) has developed different levels of eco-region maps for environmental resource management over the last decade. Levels I and II in coarse scales were replaced by level III with 84 eco-regions for the continental US, and the release of level IV is expected soon.

This study conducted comparative analyses of the NLCD and GLCC over agricultural eco-regions in the conterminous US. Although the GLCC and NLCD were developed based on different satellite sensors with different spectral and spatial resolutions, both data sets were expected to provide similar land-cover information over the continental US. The NLCD distributions within the 1 km × 1 km GLCC pixels were investigated to assess the spatial agreement and relationship of agricultural areas between the two data sets. The results have potential contributions to the knowledge of cropland correlation between fine-resolution and coarse-resolution data sets. Investigating the degree and areas of agreement between the two data sets can provide background information for interchanging less-detailed for detailed agricultural land-cover in a large area or whenever appropriate.

2. Background

The NLCD and GLCC data sets were produced based on an entire year of satellite images. The 21 classes for NLCD are modified from the well-established Anderson land-use and land-cover classification system (Anderson *et al*. 1976) (table 1). The GLCC data sets were provided in seven different land-cover classification schemes, and the International Geosphere Biosphere Program (IGBP) classification scheme in 17 classes was used for this study (table 1). Both land-cover data sets are readily available at the US Geological Survey (USGS) web site.

The positional accuracy of each pixel is important, since this study investigated the NLCD distribution within 1 km × 1 km GLCC pixels. Each LANDSAT TM image used to create the NLCD was terrain-corrected using digital terrain elevation data and geo-registered using ground control points, resulting in a root mean

Table 1. The Global Land-Cover Characteristics (GLCC) data based on the classification scheme of International Geosphere Biosphere Program (IGBP) and the National Land-Cover Data (NLCD) based on the modified Anderson land-use and land-cover classification system. The numbers in parentheses represent land-cover codes.

Land-cover types	GLCC classification scheme	NLCD classification scheme
Forested upland	Evergreen needle leaf forest (1) *Evergreen broadleaf forest (2) Deciduous needle leaf forest (3) Deciduous broadleaf forest (4) Mixed forest (5)	Evergreen forest (42) Deciduous forest (41) Mixed forest (43)
Shrubland	*Closed shrubland (6) Open shrubland (7)	Shrubland (51)
Savanna	Woody savanna (8) Savanna (9)	
Herbaceous upland	Grasslands (10)	Grasslands/herbaceous (71)
Wetland	Permanent Wetlands (11)	Woody wetlands (91) Emergent/herbaceous wetlands (92)
Herbaceous planted/cultivated	Cropland (12)	Pasture/hay (81) Row crops (82) Small grains (83) Fallow (84) Urban/recreational grasses (85)
Developed	Urban/built-up (13)	Low intensity residential (21) High intensity residential (22) Commercial/industrial/transportation (23)
Barren	*Barren lands or sparsely vegetated (16)	Bare rock (31) Quarries/mines (32) Transitional (33)
Water	*Snow/ice (15) Water (17)	*Perennial ice/snow (12) Open water (11)
Non-natural woody		Orchards/vineyard (61)
Cropland/natural vegetation mosaic	Cropland/natural vegetation mosaic (14)	

*None in study area.

square registration error of less than one pixel (30 m) (Vogelmann *et al.* 2001). In our knowledge, the registration accuracy of GLCC data has not been officially published yet. The USGS documentation stated that the goal of positional accuracy is 1 km or less for the AVHRR images used to create the GLCC. The ground control points and hydrographical features were used for geometric registration. Digital elevation data were incorporated to reduce registration errors (US Geological Survey 2005).

The NLCD was provided on a state-by-state basis (Vogelmann *et al.* 2001). The TM multi-band data were processed using an unsupervised clustering algorithm (Kelly and White 1993). The resulting spectral clusters were resolved into one of 21 thematic classes using aerial photography and ground observations. Clusters that represented more than one land-cover category were identified using various ancillary data sets and models were developed to split the confused clusters into the correct land-cover categories.

The GLCC data sets were created based on the concept of seasonal land-cover regions to present the temporal and spatial patterns of vegetation. The AVHRR Normalised Difference Vegetation Index (NDVI) composites and other ancillary geographic data, including digital elevation model, eco-region data and country or regional-level vegetation and land-cover maps, were used to classify global land-cover data. Unsupervised clustering method was applied to multi-temporal NDVI composites, and each land-cover cluster was classified based on ancillary data and expert interpretation. Post-classification was implemented using multi-source geographic data to refine classes containing two or more disparate land-cover types (Running *et al.* 1995).

Accuracy assessment of land-cover maps derived from satellite data has been an important concern of the remote sensing community (Latifovic and Olthof 2004). The 1992 NLCD has accuracy ranging from 37% (central US) to 69% (western coast) (US Geological Survey 2003). Much of the classification error occurred among the NLCD classes that aggregate into a single Anderson level I class. For example, pasture/hay was often confused with row crops, and mixed forest with deciduous forest and evergreen forest. Other sources of disagreement were between the forest and the agricultural classes and between the woody wetland class and the forest classes (Vogelmann *et al.* 2001, Yang *et al.* 2001).

For the GLCC data sets, the averaged classification accuracy was 59.4%. The highest individual class accuracies occurred in the classes of evergreen broadleaf forests (78%) and barren lands (95%). The wetlands had the lowest accuracy about 33%. The class of cropland had accuracy greater than 60%. Most errors occurred when shrubland was identified as wetlands and croplands as deciduous forest. Other errors included confusing deciduous forest with wetlands, and evergreen and deciduous forest with croplands (Scepan 1999).

3. Methods

A total of 11 agricultural-related eco-regions were selected for this study. They are distributed across the USA from coast to coast, and several of them were located in the major agricultural regions of central USA (figure 1). Subsets of the GLCC and NLCD coverage were prepared for each of the 11 eco-regions. The NLCD subsets were re-sampled from 30 m to 25 m resolution. The percentage of each land-cover in the GLCC and NLCD was computed for each studied eco-region. Several GLCC classes were omitted from this study due to non-existence in the study sites including

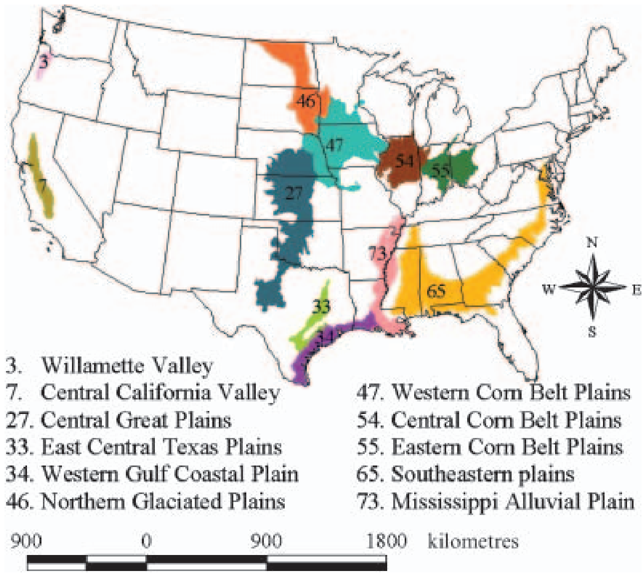

Figure 1. The locations of 11 studied eco-regions over the Continental United States.

deciduous needle leaf forest, closed shrubland, snow/ice and barren lands or sparsely vegetated.

A total of 1600 (40×40) small pixels in 25 m resolution constituted one large pixel in 1 km resolution. The land-cover distribution of NLCD within each GLCC class in the correspondent eco-region was calculated in percentage. Only the large pixels completely covered by 1600 small pixels were used for this study. Most edge pixels along each eco-region boundary were discarded due to incomplete information. A database was established to store the land-cover data of GLCC and NLCD for each studied eco-region. Each record of the database represented one 1 km $\times$ 1 km pixel, and consisted of the location of the pixel, the land-cover type of GLCC and the geo-correspondent land-cover composition of NLCD in percentage. The average and standard error for each geo-correspondent NLCD class were calculated for each major GLCC class.

4. Results

Eco-regions were identified based on the biotic and abiotic compositions of geology, physiography, vegetation, climate, soils, land-use, wildlife and hydrology. The 11 studied eco-regions were sorted into four categories mainly based on the land-use types of EPA documents (US Environmental Protection Agency 2005). The four categories were corn belt in the central US including eco-regions of 47, 54 and 55, cropland for much of the eco-regions of 7, 27, 34 and 73, a mosaic of agriculture, forest and grassland for the eco-regions of 3, 46 and 65, and a high density of pasture for the eco-region of 33.

4.1 *Corn belt in the central US*

The EPA distinguishing characteristics for level III eco-regions described that extensive agricultural cropland occupied the eco-regions of 47, 54 and 55 to produce corn and soybeans, and much of the remainder was in forage for livestock (US Environmental Protection Agency 2005). The NLCD distributions exhibited that row crops occupied more than 60% of the area, and pasture/hay was the second important land-cover type in the regions (table 2). The GLCC data provided comparable land-cover information that more than 70% of the regions were dominated by cropland, and cropland/natural vegetation mosaic covered most of the remaining area (table 3). Both NLCD and GLCC showed that forest and grassland scattered between the agricultural fields.

The NLCD distributions within the two major GLCC classes (cropland and cropland/natural vegetation mosaic) were analysed to assess the relationship between the two data sets (table 4). Our results showed that the GLCC cropland in the three eco-regions (47, 54 and 55) had a high agreement ranging from 71% to 80% to the NLCD row crops (table 4). The correspondence between GLCC cropland and NLCD pasture/hay was about 13% on average. Another analysis showed the GLCC cropland/natural vegetation mosaic corresponded to both NLCD row crops and pasture/hay in the area (table 4). The GLCC cropland was slightly related to the NLCD class of deciduous forest, and occasionally related to the NLCD grassland (table 4). Similar relationships were observed for the GLCC cropland/natural vegetation mosaic to the NLCD deciduous forest and grassland (table 4). Most standard errors were less than 0.2% due to nearly homogeneous land-cover distributions in the area and large sample sizes (table 4).

The GLCC mixed forest was a minor land-cover type for the eco-region 54 (table 3), and our results showed an extremely low spatial agreement (less than 0.33%) for the mixed forest between the GLCC and NLCD (table 4). Another low agreement (15.63%) was found for the minor GLCC class of grassland in eco-region 47. However, the GLCC deciduous broad leaf forest in eco-region 55 exhibited a reasonable agreement (50%) to the NLCD deciduous forest. For the non-vegetated urban area, our results exhibited the agreements for eco-regions 54 and 55 were greater than 60% between the two data sets.

4.2 *Cropland for much of the area*

The EPA documentation recorded that much of theses eco-regions including 7, 27, 34 and 73 were dominated by cropland (US Environmental Protection Agency 2005). However, the NLCD distributions showed these regions have relatively high land-cover diversity. Although agriculture is the major land-use type for the regions, other land-cover types such as shrubland, grassland, wetland and orchards also considerably occupied the area at significant level (table 2). The NLCD row crops, small grain and pasture/hay were equally distributed in agricultural land for the regions except for eco-region 73. The NLCD row crops occupied half of the area of eco-region 73, and both woody and herbaceous wetlands covered a quarter of the area. The GLCC distributions were less diverse than the NLCD for the regions. All four eco-regions were primarily occupied by the GLCC cropland and cropland/natural vegetation mosaic. The GLCC savanna, grassland and deciduous broadleaf forest filled up the remaining areas (table 3). The GLCC data showed that less than 1% wetland in eco-region 73, which contradicted with the NLCD wetland distribution. Similar contradictions occurred for the class of shrubland in eco-region 34. Both GLCC and NLCD showed similar percentages of grassland distribution for eco-region 27.

Three major GLCC classes for each eco-region were studied to analyse their spatial relationship with the NLCD classes (table 5). The GLCC cropland did not merely correspond to the NLCD row crops, small grains and pasture/hay, but the

Table 2. The land-cover distribution in percentage (%) of National Land-Cover Data (NLCD) for each studied eco-region.

Eco-region / Classes	Corn belt			Cropland				Mosaic of agriculture and other vegetation			Pasture
	47	54	55	7	27	34	73	3	46	65	33
Water	0.98	1.10	0.71	1.03	1.04	4.71	7.56	1.98	1.60	1.40	1.91
Low intensity residential	0.80	1.65	3.01	3.15	0.44	1.96	0.82	5.07	0.19	1.11	0.36
High intensity residential	0.25	1.97	0.49	0.15	0.18	1.27	0.28	0.04	0.04	0.21	0.15
Commercial/industrial/ transportation	1.19	1.28	1.28	1.45	0.39	1.66	0.52	2.62	0.32	0.61	0.68
Bare rock	0.01	0.02	<0.01	1.48	0.32	0.53	0.07	0.10	0.07	0.04	0.37
Quarries/mines	0.04	0.07	0.05	0.03	0.01	0.05	0.01	0.06	0.02	0.11	0.20
Transitional	<0.01	<0.01	0.06	0.01	0.01	0.04	0.09	0.45	<0.01	3.91	<0.01
Deciduous forest	4.56	8.24	11.76	0.3	1.74	6.39	2.38	5.78	1.86	17.24	22.75
Evergreen forest	0.03	0.35	0.31	0.25	0.66	5.24	0.46	18.99	0.01	21.25	5.94
Mixed forest	0.16	0.06	0.04	0.38	0.12	1.35	1.29	9.68	0.01	15.59	2.49
Shrubland	0.03	0.01	<0.01	1.46	9.25	11.22	<0.01	1.44	0.22	0.02	10.89
Orchards/vineyard	0	0.01	0	18.76	<0.01	0.06	0	3.55	0	0.02	<0.01
Grasslands/herbaceous	5.18	0.63	0.04	19.02	35.75	8.23	0.25	3.00	9.90	0.29	8.80
Pasture/hay	14.75	12.16	18.50	16.47	7.94	16.41	5.69	35.24	19.32	10.05	36.82
Row crops	68.77	69.76	62.10	20.44	22.04	17.04	48.71	5.65	49.34	17.29	5.77
Small grains	1.05	0.13	<0.01	13.89	19.30	10.62	5.11	5.10	11.38	<0.01	0.15
Fallow	<0.01	0	0	0.02	0.09	0	0	0.10	1.05	0	<0.01
Urban/recreation grasses	0.34	1.14	0.83	0.21	0.11	0.59	0.38	0.49	0.08	0.24	0.08
Woody wetlands	0.82	1.17	0.72	0.23	0.09	2.01	15.41	0.23	0.11	10.19	1.37
Emergent/herbaceous wetlands	1.03	0.24	0.09	1.27	0.53	10.64	10.97	0.43	4.47	0.43	1.26

Table 3. The land-cover distribution in percentage (%) of Global Land-Cover Characteristics (GLCC) for each studied eco-region.

Eco-region	Corn belt			Cropland				Mosaic of agriculture and other vegetation			Pasture
Classes	47	54	55	7	27	34	73	3	46	65	33
Evergreen needle leaf forest	0.02	0.34	0.08	0.89	0.57	4.79	3.43	59.83	0.03	41.21	2.56
Evergreen broadleaf forest	0	0	0	0	<0.01	0	<0.01	0	0	<0.01	0
Deciduous broadleaf forest	0.06	0.47	3.45	0.26	0.03	0.25	6.88	0	0.74	9.57	1.42
Mixed forest	0.19	3.48	0.45	2.28	0.03	0.06	0.34	0.07	0.01	3.46	0
Open shrubland	<0.01	0.03	0	5.87	0.03	0.05	0.01	0.02	0.06	<0.01	0.01
Woody savanna	0.14	0.38	0.40	31.57	7.26	2.17	6.23	0.05	0	0.23	0.45
Savanna	0	0	0	14.25	0	0.87	0.21	0	0	0.01	0.06
Grasslands	1.79	0.38	0.14	2.58	37.40	10.87	1.53	0.29	6.95	0.13	0.60
Permanent wetlands	0	0	0	0	0	0.28	0.19	0	0	<0.01	<0.01
Cropland	93.51	77.58	72.07	35.93	27.65	55.44	45.76	33.77	54.17	24.05	72.77
Urban/built-up	0.87	3.57	3.34	2.29	0.58	2.16	0.79	5.48	0.12	1.09	0.27
Cropland/natural vegetation mosaic	3.13	13.34	19.90	3.73	25.97	17.50	28.50	0.09	37.09	19.79	20.92
Water	0.30	0.43	0.21	0.36	0.48	5.55	6.13	0.40	0.83	0.44	0.94

Table 4. The spatial correspondence of GLCC classes to the NLCD classes for the three corn belt eco-regions 47, 54 and 55. The numbers and numbers in parentheses represented the averaged NLCD composition in percentage (%) and standard error, respectively, for the GLCC classes in the studied eco-regions.

GLCC	Cropland			Cropland/natural vegetation mosaic			Grassland	Mixed forest	Deciduous broadleaf forest	Urban/built-up	
NLCD	47	54	55	47	54	55	47	54	55	54	55
Urban	1.73 (0.02)	1.20 (0.03)	1.44 (0.03)	5.10 (0.08)	8.29 (0.07)	7.16 (0.06)	3.34 (0.09)	1.76 (0.03)	1.92 (0.05)	64.69 (0.34)	60.8 (0.34)
Deciduous forest	4.15 (0.02)	5.30 (0.04)	7.48 (0.04)	10.74 (0.15)	17.98 (0.14)	19.61 (0.12)	9.52 (0.20)	28.10 (0.21)	49.99 (0.42)	8.18 (0.19)	11.08 (0.23)
Mixed forest	0.14 (0.00)	0.03 (0.00)	0.02 (0.00)	0.76 (0.02)	0.15 (0.01)	0.07 (0.00)	0.36 (0.02)	0.33 (0.01)	0.33 (0.02)	0.04 (0.01)	0.10 (0.01)
Grassland	4.95 (0.02)	0.51 (0.01)	0.04 (0.00)	6.62 (0.13)	1.09 (0.04)	0.02 (0.00)	15.63 (0.26)	0.84 (0.02)	0.06 (0.01)	0.80 (0.05)	0.02 (0.00)
Pasture/hay	14.00 (0.03)	10.38 (0.04)	14.81 (0.05)	31.62 (0.23)	23.04 (0.14)	32.28 (0.15)	25.32 (0.29)	18.27 (0.14)	26.28 (0.35)	3.46 (0.13)	6.90 (0.19)
Row crops	71.20 (0.05)	80.16 (0.08)	74.43 (0.08)	37.52 (0.28)	41.34 (0.23)	37.04 (0.17)	34.08 (0.38)	43.26 (0.28)	15.18 (0.26)	9.34 (0.36)	12.70 (0.37)
Small grains	0.92 (0.01)	0.11 (0.00)	0.00 (0.00)	2.51 (0.06)	0.23 (0.01)	0.00 (0.00)	4.47 (0.12)	0.18 (0.01)	0.00 (0.00)	0.06 (0.02)	0.00 (0.00)
Others	3.05	2.34	1.80	5.89	8.03	3.89	7.64	7.59	6.57	13.43	8.40

NLCD orchard (eco-region 7), grassland (eco-region 27) and wetland (eco-regions 34 and 73). The GLCC cropland/natural vegetation mosaic was mainly related to the NLCD grassland (50.45%) in eco-region 27 and row crops (59.59%) in eco-region 73. For the eco-region 34, the GLCC cropland/natural vegetation mosaic evenly corresponded to multiple NLCD classes including row crops (26.76%), shrubland (23.87%) and pasture/hay (18.59%). The NLCD classification scheme does not have the classes of savanna and woody savanna. The results for eco-region 7 showed that the GLCC woody savanna primarily corresponded to the NLCD orchard (32.04%), and the GLCC savanna to the NLCD grassland (59.20%). The grassland between both data sets had agreements of 32.86% for eco-region 27 and 9.02% for eco-region 34 (table 5). The GLCC grassland corresponded to the NLCD shrubland for both eco-regions of 27 and 34, and to the NLCD row crops for eco-region 34. The GLCC deciduous broadleaf forest had a low agreement with the NLCD deciduous forest (3.27%) for eco-region 73, but was strongly related to the NLCD woody wetland (65.40%).

4.3 *Mosaics of agriculture, forest and grassland*

According to the documentation for EPA Level III eco-region, the eco-regions of 3, 46 and 65 were characterised by agricultural cropland and natural vegetation including forest (eco-regions 3 and 65), grassland (eco-region 46) and woodland (eco-region 65) (US Environmental Protection Agency 2005). Both NLCD and GLCC distributions were similar to the EPA documentation for the regions (tables 2 and 3).

Three major GLCC classes for each eco-region were studied to analyse their spatial relationship with the NLCD classes. These were cropland (eco-regions 3, 46 and 65), cropland/natural vegetation mosaic (eco-regions 46 and 65), evergreen needle leaf forest (eco-regions 3 and 65), grassland (eco-region 46) and urban (eco-region 3). The two GLCC classes of cropland and cropland/natural vegetation mosaic as usual corresponded to the NLCD pasture/hay, row crops and small grains in the regions, and were related to the NLCD forest for eco-region 65 (figures 2a and 2b). The GLCC evergreen needle leaf forest was related to all types of NLCD forests in eco-regions 3 and 65 (figure 2c). The GLCC grassland in eco-region 46 had an agreement of 23% with the NLCD grassland, and was equally related to the NLCD pasture/hay, row crops and small grains (figure 2d). Urban area occupied only 5.5% of the area of eco-region 3, and the GLCC urban had a high agreement of 63% with the NLCD urban.

4.4 *High density of pasture*

The eco-region 33 was mainly covered by pasture and range according to the EPA documentation (US Environmental Protection Agency 2005). The NLCD distributions did not exhibit homogeneous land-cover information, but showed that the pasture/hay occupied about 36% of the area, and deciduous forest and shrubland were the other two major land-cover types for the region (table 2). The GLCC exhibited that more than 90% of the region was dominated by either cropland or cropland/natural vegetation mosaic (table 3).

The GLCC cropland and cropland/natural vegetation mosaic were mainly correlated to the NLCD agricultural class of pasture/hay only, since the NLCD row crops and small grains were minor classes for the region (figure 3). The GLCC cropland was particularly related to the NLCD deciduous forest (24.83%), and the GLCC cropland/natural vegetation mosaic was equally related to the NLCD deciduous forest (15.12%), shrubland (17.95%) and grassland (13.86%).

Table 5. The spatial correspondence of GLCC classes to the NLCD classes for the four eco-regions 7, 27, 34 and 73. The numbers and numbers in parentheses represented the averaged NLCD composition in percentage (%) and standard error, respectively, for the GLCC classes in the studied eco-regions.

GLCC	Cropland				Cropland/natural vegetation mosaic			Woody savanna	Savann-a	Grassland		Deciduous broadleaf forest
NLCD	7	27	34	73	27	34	73	7	7	27	34	73
Deciduous forest	0.07 (0.00)	2.25 (0.02)	8.67 (0.07)	1.96 (0.02)	1.60 (0.02)	3.54 (0.07)	2.41 (0.03)	0.43 (0.02)	0.73 (0.04)	1.02 (0.01)	1.01 (0.04)	3.27 (0.11)
Evergreen forest	0.14 (0.01)	0.24 (0.00)	6.18 (0.08)	0.29 (0.01)	0.62 (0.02)	3.43 (0.09)	0.47 (0.01)	0.26 (0.01)	0.41 (0.02)	0.85 (0.01)	1.06 (0.03)	0.47 (0.03)
Shrubland	0.46 (0.01)	0.58 (0.01)	9.33 (0.10)	0.00 (0.00)	4.24 (0.05)	23.87 (0.26)	0.00 (0.00)	0.89 (0.03)	3.07 (0.09)	20.30 (0.09)	12.53 (0.23)	0.00 (0.00)
Orchard	18.66 (0.24)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.14 (0.02)	0.00 (0.00)	32.04 (0.29)	7.68 (0.24)	0.00 (0.00)	0.29 (0.04)	0.00 (0.00)
Grassland	4.74 (0.10)	26.04 (0.10)	8.22 (0.08)	0.41 (0.02)	50.45 (0.11)	12.06 (0.18)	0.04 (0.01)	15.20 (0.22)	59.20 (0.47)	32.86 (0.08)	9.02 (0.24)	0.02 (0.00)
Pasture/hay	19.09 (0.18)	9.01 (0.04)	19.51 (0.12)	5.70 (0.05)	6.15 (0.03)	18.59 (0.19)	7.54 (0.06)	17.91 (0.18)	11.51 (0.26)	7.60 (0.04)	9.26 (0.17)	2.22 (0.07)
Row crops	29.73 (0.25)	37.63 (0.12)	11.03 (0.11)	58.28 (0.16)	14.72 (0.07)	26.76 (0.31)	59.59 (0.16)	16.35 (0.20)	4.73 (0.17)	17.26 (0.08)	49.34 (0.48)	15.28 (0.31)
Small grains	20.86 (0.24)	21.42 (0.11)	16.42 (0.16)	5.34 (0.04)	20.19 (0.08)	1.98 (0.11)	7.69 (0.07)	10.66 (0.16)	3.16 (0.14)	16.89 (0.07)	0.30 (0.05)	1.87 (0.08)
Woody wetland	0.29 (0.01)	0.20 (0.06)	2.16 (0.04)	7.27 (0.07)	0.07 (0.00)	0.40 (0.03)	15.48 (0.13)	0.25 (0.02)	0.07 (0.01)	0.02 (0.00)	0.13 (0.02)	65.40 (0.42)
Herbaceous wetland	1.77 (0.07)	0.88 (0.02)	10.97 (0.13)	13.99 (0.13)	0.41 (0.01)	3.46 (0.12)	0.49 (0.03)	0.97 (0.05)	0.63 (0.05)	0.31 (0.01)	4.61 (0.18)	4.69 (0.19)
Others	4.19	1.75	7.51	6.76	1.55	5.77	6.29	5.04	8.81	2.89	12.45	6.78

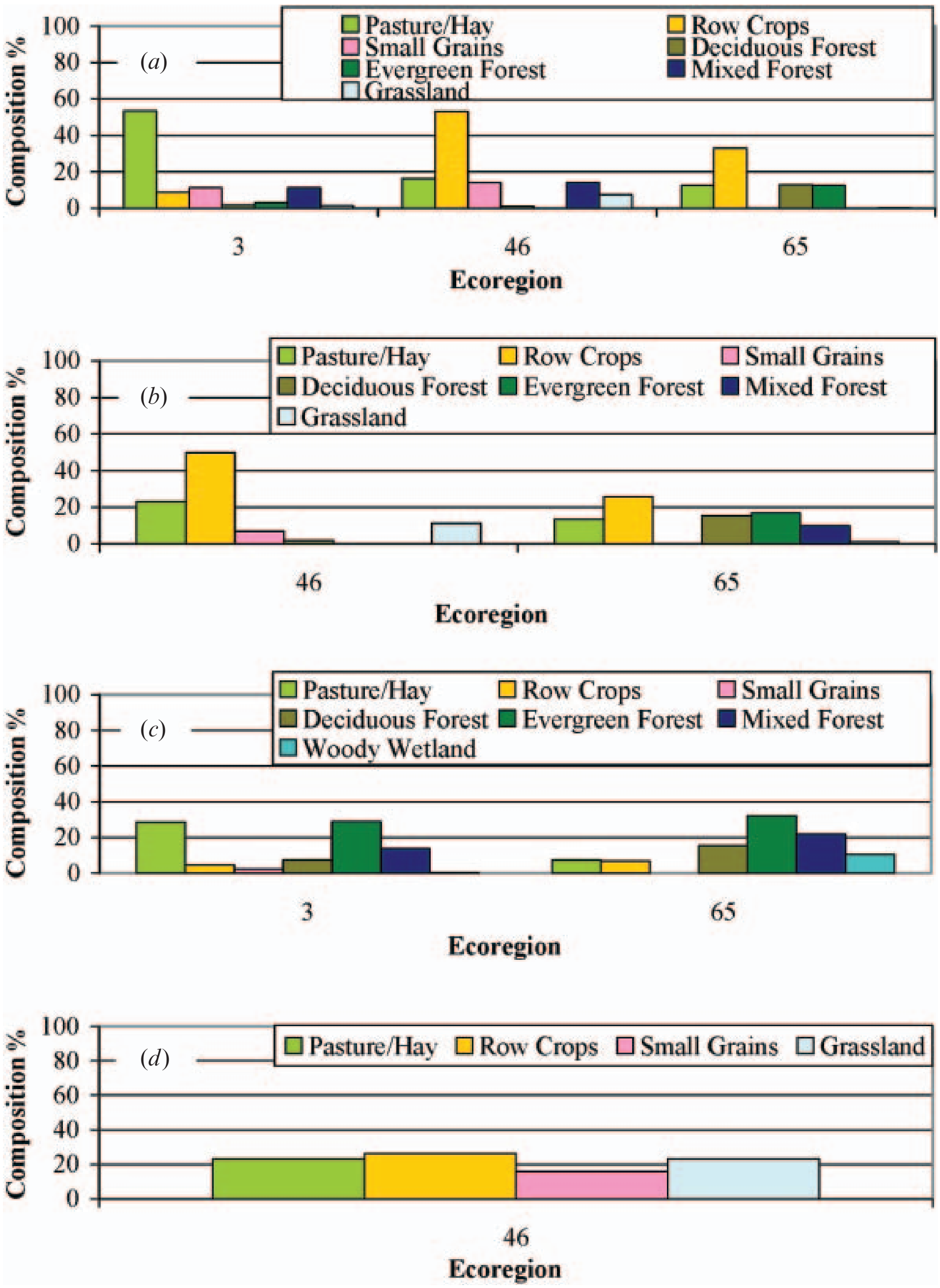

Figure 2. The spatial correspondence of GLCC (*a*) cropland, (*b*) cropland/natural vegetation mosaic, (*c*) Evergreen needle leaf forest and (*d*) grassland to the related NLCD classes for the eco-regions 3, 46 and 65.

5. Discussion

The EPA documentation of land-use types agreed to the NLCD distributions for most study sites but the eco-regions of Central California Valley (7), Central Great Plains (27), East Central Texas Plains (33) and Western Gulf Coastal Plain (34).

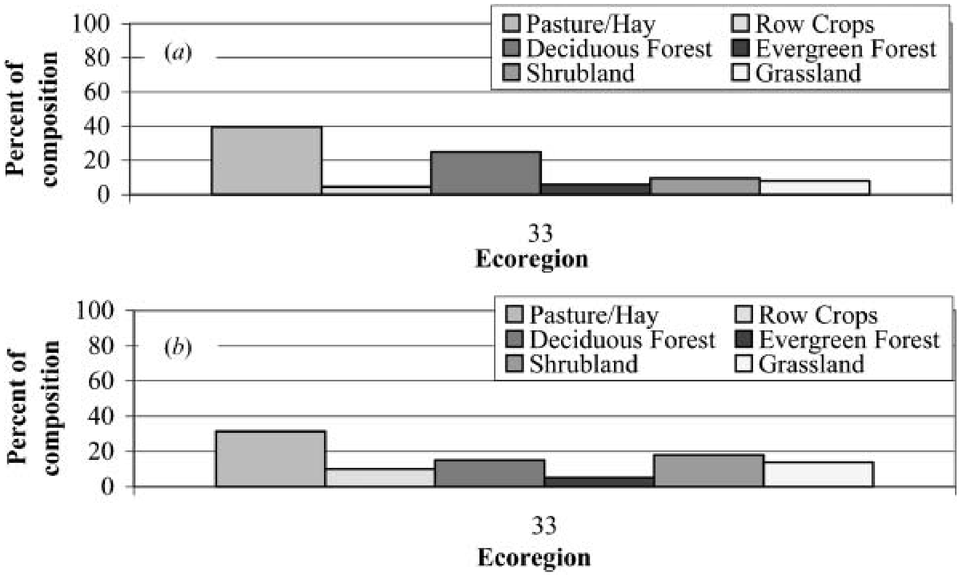

Figure 3. The spatial correspondence of GLCC (*a*) cropland and (*b*) cropland/natural vegetation mosaic to the related NLCD classes for the eco-region 33.

Similar data contradictions between the EPA eco-region characterisations and GLCC distributions occurred in the eco-regions of 7 and 27. Both NLCD and GLCC data sets had comparable numerical distributions for agricultural lands except for the eco-regions of 7, 33 and 34. The entire comparisons revealed that the NLCD classification produced less croplands compared to the EPA documentation and GLCC data sets, which agree to the findings by Wardlow and Egbert (2003) that NLCD appeared to underestimate cropland for the state of Kansas.

Our study found in general that the GLCC cropland primarily referred to the NLCD row crops, pasture/hay and/or small grains, though it corresponded to grasslands, deciduous and evergreen forest depending on the NLCD distributions. The accuracy assessment indicated that NLCD pasture/hay was often misclassified as row crops or vice versa (Vogelmann *et al.* 2001). Moreover, the NLCD pasture/hay was confused with the grassland (Wardlow and Egbert 2003). The NLCD pasture/hay class was defined as grass, legumes, or grass legume mixtures planted for livestock grazing or the production of seed or hay crops (US Geological Survey 2003), which was difficult to distinguish from grassland or cropland in 30 m resolution. Other misclassifications occurred between the NLCD agricultural classes and forest classes (Vogelmann *et al.* 2001). Overall, the NLCD classification generated more grassland and forestland than the GLCC data sets over the studied eco-regions.

The GLCC class of cropland/natural vegetation mosaic carries information of croplands and other vegetated land-cover due to the mixtures of plants in 1 km unit. The applications of this GLCC class are challenging when analysing the geophysical properties of this land-cover for soil and water modelling or using the land-cover information for interaction studies between vegetation and atmosphere. Our results demonstrated that the GLCC cropland/natural vegetation mosaic mainly corresponded to the NLCD row crops and pasture/hay, secondly to the NLCD grassland and forest, and occasionally to the NLCD small grains and shrubland, which provide valuable insights to data users. However, the mosaic class is a disadvantage of the frequent-updated global land-cover data sets.

The GLCC woody savanna and savanna were not included in the NLCD classification schemes. Although the class of woody savanna and savanna existed merely in few eco-regions, the spatial correspondence for the two unique GLCC classes concerns data users. The results for eco-region 7 showed that the GLCC woody savanna corresponded to the NLCD orchard. Other eco-regions exhibited that GLCC woody savanna was spatially related to the NLCD grassland (eco-region 27), small grains (eco-regions 27 and 34), row crops (eco-region 73), woody wetland (eco-region 73) and herbaceous wetlands (eco-region 34). The correspondences for GLCC woody savanna were inconsistent, since no common features were found between the corresponded NLCD classes besides greenness. The GLCC savanna corresponded to the NLCD grassland in eco-region 7 and to the NLCD row crops and pasture/hay in eco-region 34. Our assumptions are that the visual texture and spectral reflectance of savanna must be similar to that of grassland and pasture/hay.

The EPA characterisations for level III eco-regions do not report any wetland distribution for eco-regions 34, 65 and 73, though the three eco-regions were near either the Atlantic Ocean or the Gulf of Mexico. The GLCC data sets revealed less than 1% wetlands for each of the three eco-regions, while the NLCD distributions exhibited more than 10% wetland for the same areas. Most of eco-region 73 is located in the state of Florida, and the investigation by Clark (2004) reported that wetlands cover almost 30% of Florida, which was close to the 26% of NLCD. It is evident that GLCC data underestimated wetlands when compared to the NLCD. Additional studies are required to accurately classify wetland distributions for global land-cover data sets.

Although urban/built-up occupied a small proportion of area for most eco-regions, cropland has been shrinking over recent years partially due to urbanisation. The numerical proportions of urban distributions of GLCC and NLCD were close for most eco-regions. Our results demonstrated that a reasonable agreement between the GLCC and NLCD ranging from 39% to 65% for the urban class. The GLCC urban/built-up primarily corresponded to the NLCD low intensity residual, and secondly to the NLCD high intensity residential and commercial/industrial/transportation.

No sampling technique was involved in this study, since each pixel was counted for analyzing data agreements and relationships between the GLCC and NLCD. The total number of 1 km $\times$ 1 km pixel exceed 100,000 for all eco-regions except for the smallest eco-region of 3. The largest eco-region of 65 holds more than 900,000 1 km unit pixels. The huge sizes of samples consequently yield negligible standard errors, which indicated the results of this study are reliable and ready for further applications.

6. Conclusions

This land-cover study has been focused on agricultural eco-regions, because agricultural distributions and their associated management practices have greater impacts on soil and water conservation compared to other vegetated lands. Each eco-region is expected to represent areas in similar or homogeneous land-cover/land-use. The NLCD distribution within each major GLCC class in 1-km unit was analysed for each of the 11 agricultural-related eco-regions. Each pixel was counted for this study, and the huge sample sizes resulted in small standard errors. The GLCC classes of cropland and cropland/natural vegetation mosaics in this study primarily corresponded to the NLCD row crops, pasture/hay and small grains, and occasionally related to the NLCD grassland and forests. The eco-regions located in the corn belt plains were dominated by homogeneous croplands according to the NLCD distributions, and the GLCC cropland mostly represented the NLCD row crops. Diverse NLCD distributions were observed in the remaining eco-regions; therefore, the data correspondences for the GLCC cropland and cropland/natural vegetation mosaics were complicated. The GLCC savanna was mainly related to the NLCD grassland and pasture/hay due to similar spectral reflectance and visual texture. However, the correspondence of GLCC woody savanna was variable depending on the NLCD distributions. Both land-cover data sets had reasonable agreements for the urban areas, which permit the frequent-updated global land-cover data sets to be used for monitoring urban growth.

Satellite data provide land-cover/land-use information for research studies. The land-use is dynamic, and the need of multi-temporal land-cover data is increasing. Several land-cover data sets based on coarse-resolution satellite images have been produced over the last few years. The coarse-resolution data provided less detailed land-cover information, and have been restricted for global or national studies so far. However, the currently available detail-oriented land-cover data was produced in 1992, which require a critical update. Near real-time studies certainly need to rely on the global land-cover data sets. This study provided information of the degree of spatial agreement and relationship between the GLCC and NLCD data sets within 11 eco-regions. Other data sources such as local land-cover maps or field survey are applicable to verify the GLCC distributions for the areas in low agreement. For those areas the GLCC data showed reasonable agreement and relationship to the NLCD, the recently updated global land-cover data could serve as an input for regional studies to enhance near real-time environmental research.

Acknowledgements

The authors would like to thank the USDA-Agricultural Research Service (ARS) in Temple, TX for supporting this research through the Specific Cooperate Agreement No. 58-6206-1-005.

References

ANDERSON, J.R., HARDY, E.E., ROACH, J.T. and WITMER, R.E., 1976, A land use and land cover classification system for use with remote sensor data. U.S. Geological Survey, Reston, VA, U.S. Geological Survey Professional Paper 964, 28 pp.

ARNOLD, J.G., SRINIVASAN, R., MUTTIAH, R.S. and WILLIAMS, J.R., 1998, Large area hydrologic modeling and assessment: Part I, Model development. *Journal of the American Water Resources Association*, **34**, pp. 73–89.

BROWN, J.F., LOVELAND, T.R., MERCHANT, J.W., REED, B.C. and OHLEN, D.O., 1993, Using multisource data in global land cover characteristics: Concepts, requirements and methods. *Photogrammetric Engineering and Remote Sensing*, **59**, pp. 977–987.

CLARK, M., 2004, Florida wetlands: Extension web site. Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL, Fact Sheet SL217, 2 pp.

EIDENSHINK, J.C. and FAUNDEEN, J.L., 1994, The 1 km AVHRR global land dataset-first stages in implementation. *International Journal of Remote Sensing*, **15**, pp. 3443–3462.

FRIEDL, M.A., MCLVER, D.K., HODGES, J.C.F., ZHANG, X.Y., MUCHONEY, D., STRAHLER, A.H., WOODCOCK, C.E., GOPAL, S., COOPER, A., BACCINI, A., GAO, F.

and SCHAAF, C., 2002, Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment*, **83**, pp. 287–302.

GIRI, C., ZHU, Z.L. and REED, B., 2005, A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. *Remote Sensing of Environment*, **94**, pp. 123–132.

HOMER, C.G., HUANG, C., YANG, L. and WYLIE, B., 2002, Development of a Circa 2000 Landcover Database for the United States. In *ASPRS 2002 Annual Conference*, 22–26 April 2002, Washington, DC.

KELLY, P.M. and WHITE, J.M., 1993, Preprocessing remotely sensed data for efficient analysis and classification. In *SPIE-Applications of Artificial Intelligence 1993: Knowledge-Based Systems in Aerospace and Industry*, April 1993, pp. 24–30.

LATIFOVIC, R. and OLTHOF, I., 2004, Accuracy assessment using sub-pixel fractional error matrices of global land-cover products derived from satellite data. *Remote Sensing of Environment*, **90**, pp. 153–165.

LOVELAND, T.R. and MERCHANT, J.M., 2004, Ecoregions and ecoregionalisation: Geographical and ecological perspectives. *Environmental Management*, **34**(Supplement 1), pp. S1–S13.

LOVELAND, T.R., REED, B.C., BROWN, J.F., OHLEN, D.O., ZHU, Z., YANG, L. and MERCHANT, J.W., 2000, Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, **21**, pp. 1303–1330.

OMERNIK, J.M. and BAILEY, R.G., 1997, Distinguishing between watersheds and ecoregions. *Journal of the American Water Resources Association*, **33**, pp. 935–949.

PLUMMER, S.E., 2000, Perspectives on combining ecological process models and remotely sensed data. *Ecological Modelling*, **129**, pp. 169–186.

RUNNING, S.W., LOVELAND, T.R., PIERCE, L.L., NEMANI, R.R. and HUNT, E.R., 1995, A remote sensing based vegetation classification logic for global land cover analysis. *Remote Sensing of Environment*, **51**, pp. 39–48.

SCEPAN, J., 1999, Thematic validation of high-resolution global land-cover datasets. *Photogrammetric Engineering and Remote Sensing*, **65**, pp. 1051–1060.

SMITH, R.A., SCHWARZ, G.E. and ALEXANDER, R.B., 1997, Regional interpretation of water-quality monitoring data. *Water Resources Research*, **33**, pp. 2781–2798.

US ENVIRONMENTAL PROTECTION AGENCY 2005, Level III Ecoregions—Ecoregion descriptions (US level III descriptions). Available on line at: http://www.epa.gov/wed/pages/ecoregions/level_iii.htm (accessed 4 August 2005).

US GEOLOGICAL SURVEY 2003, National land cover class characterisation, USGS EROS Data Center, Sioux Falls, South Dakota. Available on line at: http://landcover.-usgs.gov/natllandcover.asp (accessed 20 October 2003).

US GEOLOGICAL SURVEY 2005, AVHRR 1-km global land 10-day composites project/campaign document. Available on line at: http://eosims.cr.usgs.gov:5725/CAMPAIGN_DOCS/avhrr_gc_proj_camp.html (accessed 6 April 2005).

VOGELMANN, J.E., SOHL, T. and HOWARD, S.M., 1998, Regional characterisation of land cover using multiple sources of data. *Photogrammetric Engineering and Remote Sensing*, **61**, pp. 45–57.

VOGELMANN, J.E., HOWARD, S.M., YANG, L., LARSON, C.R., WYLIE, B.K. and VAN DRIEL, N., 2001, Completion of the 1990s national land cover dataset for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. *Photogrammetric Engineering and Remote Sensing*, **67**, pp. 650–662.

WARDLOW, B.D. and EGBERT, S.L., 2003, A state-level comparative analysis of the GAP and NLCD land-cover data sets. *Photogrammetric Engineering & Remote Sensing*, **69**, pp. 1387–1397.

YANG, L., STEHMAN, S.V., SMITH, J.H. and WICKHAM, J.D., 2001, Short communication: Thematic accuracy of MRLC land-cover for the eastern United States. *Remote Sensing of Environment*, **76**, pp. 418–422.